Kalmasoft

Arabic Personal Names Transcription System

MAPS Ono® Lite

(Transcription)

Version 1.20

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

User Manual

Arabic personal Names Transcription System

MAPSOno® Lite Transcription

Welcome!

Thanks for using Kalmasoft products, this document is the English version of MAPSOno® Lite Transcription version 1.20 user manual. MAPSOno® Lite is a multifunctional Arabic personal names transcription system.

In itself, MAPSOno® is one of Kalmasoft MAPS Suit components, it comes in two editions, Lite and Professional; for most of your ordinary purposes the Lite edition will be fairly enough, this manual contains few hints and references to the professional edition so if you decide to obtain one for your business we will be pleased to submit detailed information.

About Kalmasoft MAPS

MAPS (Multitasking Arabic Processing System) is our professional multilingual processing system for Arabic, a modular and compact yet versatile system capable of dealing with many tasks related to Arabic content management and NLP in general; MAPS Suit comprises MAPSOno® for personal names, MAPSOno® for place names, MAPSOrtho® for orthography, and MAPSSeman® for all Arabic semantic content management purposes; all components come in Lite and Pro editions.

This manual describes the usage of one of the products available only from Kalmasoft i.e. MAPSOno® Lite Transcription, it does not show the wide spectrum of MAPS Suit, please refer to Kalmasoft website for further information on any specific product.

Copyright Kalmasoft ® 2011 All rights reserved.

All content are subject to change without prior notice.

DOCMOLTR150311

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Contents

Coverage of this document 7
Copyright and Illegal Usage Policy 7
Organization of this manual 8
Installation instructions 9
Uninstalling the program 10
Program folder contents description 12
A quick start 13
Program interfaces 14
Master function buttons 15
Quick options pane 15
Transcription progress bar 15
Input interface 16
Buttons 16
Menus 16
Input pane 17
Single line input 17
Load file pane 18
View interface 18
Result display area 18
Master language selector menu 20
Status display area 20
Error messages and status hints 21
Settings interface 22
Main setting buttons 22
Layout direction and widgets placement 23
User interface language selector 23
Source language pane 23
Vocalization mode selection pane 23
Input file format and encoding pane 25
Output file format and encoding pane 26

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Romanization system selection pane 27
Fine Tuning Interface 31
Arabic language settings 31
Arabic language formality level 31
Target language varieties and regional settings 32
Orthographic variants settings 33
Target language writing subsystem 33
Program default settings 34
Saving user custom settings 34
Running the program 35
Setting the program for input 35
Preparing the names for input 36
Setting the maximum length of the name 36
Setting the names delimiter 36
Setting the ID prefix 37
Honorific titles translation 37
Allowing the insertion of special letters 37
Allowing extended Arabic letters 37
Modes of vocalization 38
Fuzzy vocalization 38
Partial vocalization 38
Manual vocalization 38
Custom vocalization 40
Indirect vocalization 40
Vocalizing a name using built-in database 40
Error detection and indication strategy 41
Orthographic variants 42
Displaying variants 44
Precision calculation 45
Precision bars 45
Preparing for output 46
Printing 46

Saving files in different formats and encodings 46
Output sorting and filtering 47
Transcription scenarios 49
General directions 49
A. Transcribing names for an official foreign client 49
B. Transcribing names for a foreign client 50
C. Transcribing names for an official home client 50
Using the included sample files 50
Tips and tricks 51
MAPSOno® frequently asked questions 52
Registering MAPSOno® Lite (Transcription) 52
Input interface issues 53
Can't load files 53
Program halts during load 53
Setting issues 53
My settings don't work 53
My new settings do not seem to affect the results 53
Can't save my settings 53
Display issues 54
Interface language sounds "Greek" 54
Transcribed names not showing up 54
Precision indicator bars shown scattered 54
Precision bars never hit 100% 54
I am getting too many variants 54
Can't copy items 54
Print issues 55
Printout not showing results 55
Mojibake and strange symbols when printing 55
How can I get rid of Kalmasoft logo 55
File saving issues 55
Can't save files 55
Can't find my results in the saved file 55

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Filtering and sorting issues 55
Filtering doesn't work 55
Can't sort results 56
Appendixes 57
Appendix A. Error messages and status hints 57
Appendix B. Romanization systems 59
Glossary of terms 60

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Coverage of this document

This document is the full text user manual for MAPSOno® Lite (Transcription) version 1.20.

Copyright and Illegal Usage Policy

Disclaimer of Liability

In preparation of this document, every effort has been made to offer the most current, correct, and clearly expressed information possible. Nevertheless, inadvertent errors in information may occur. In particular but without limiting anything here, Kalmasoft disclaims any responsibility for typographical errors and accuracy of the information that may be contained in this manual. The information and data included herein have been compiled by our staff from a variety of sources, and are subject to change without notice to you. Kalmasoft makes no warranties or representations whatsoever regarding the quality, content, completeness, suitability, adequacy, sequence, accuracy, or timeliness of such information and data. In any situation where the official sent documents of Kalmasoft differ from the text contained in this manual, the official documents take precedence. The information and data made available in this document are provided "as is" without warranties of any kind.

Disclaimer of reliability

Kalmasoft makes no representations or warranties regarding the condition or functionality of this software, its suitability for use, or that this will be uninterrupted or error-free.

Disclaimer of damages

By using Kalmasoft MAPSOno® Lite (Transcription), you assume all risks associated with the use of this software, including any risk to your computer, software or data being damaged by any virus, software, or any other file which might be transmitted or activated via this software. We shall not in any event be liable for any direct, indirect, punitive, special, incidental, or consequential damages, including, without limitation, lost revenues, or lost profits, arising out of or in any way connected with the use or misuse of the software or lack of information in this manual.

Disclaimer of endorsement

Kalmasoft does not favor one group over another, and any references herein to any country, organizations, specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, do not necessarily constitute or imply its endorsement or recommendation by us.

Copyright information

The graphics and contents on this manual are the copyrighted work of Kalmasoft and contain proprietary trademarks and trade names of the Company. No part of this document can be copied without a prior written consent from Kalmasoft.

Trademarks information

All software products mentioned in this document are registered trademarks of their respective holders.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Organization of this manual

This manual follows specific conventions in showing the aspects and usage of MAPSOno® Lite (Transcription), these are detailed below:

Examples:

Examples are shown on a brick red highlighted background, as shown below, target language is English unless otherwise stated.

Muhammad	محمد

Interface elements:

- Buttons are shown between parenthesis in thick red font like this: (**Button**)
- Menus, check boxes, and radio buttons are shown between brackets in green italicized font like this: box [*check text*]

Notes:

> Shown inside a box with black borders and red shadow, notes quote important information about specific subjects.

Cautions:

> cautions are shown inside a red box and display important information on how to use specific features of the program.

Keyboard shortcuts:

Shortcut	Details
Ctrl+A	Press both ctrl key and the letter A
F1	Press function key F1

References:

References are shown in blue, page number also shown inside parenthesis.

Installation instructions

MAPSOno® Lite (Transcription) comes packaged in an installable single exe file, all that you need to do is double click the file icon (OnoLiteTrans.exe) and follow the instructions on the screen.

The program works under the following Microsoft operating systems (Windows Vista, Windows XP, Windows 2000, Windows NT, Windows7); (Windows Me, Windows98, Windows95) were not tested. The program will create a subfolder "MAPSOno" in the system folder "Program Files" including folders with names (Input, Output, Samples, Documents) as well as the basic program files, the program also adds its icon to the desktop.

The program does not add any entries to the Start menu "Start" or entries to the system registry, thus it can be removed completely by deleting its folder from the file system, the program comes with an uninstaller tool for this purpose, this is described in the (Uninstalling the program) section on page (10) which shows you step by step how to remove the program from your system.

If the program is installed successfully on an operating system (e.g. Windows XP) its folder should show the contents as shown in Figure 1**Figure 2** below, which may look slightly different from what appears in the figure depending on the edition you obtained, the version you are using, and the operating system.

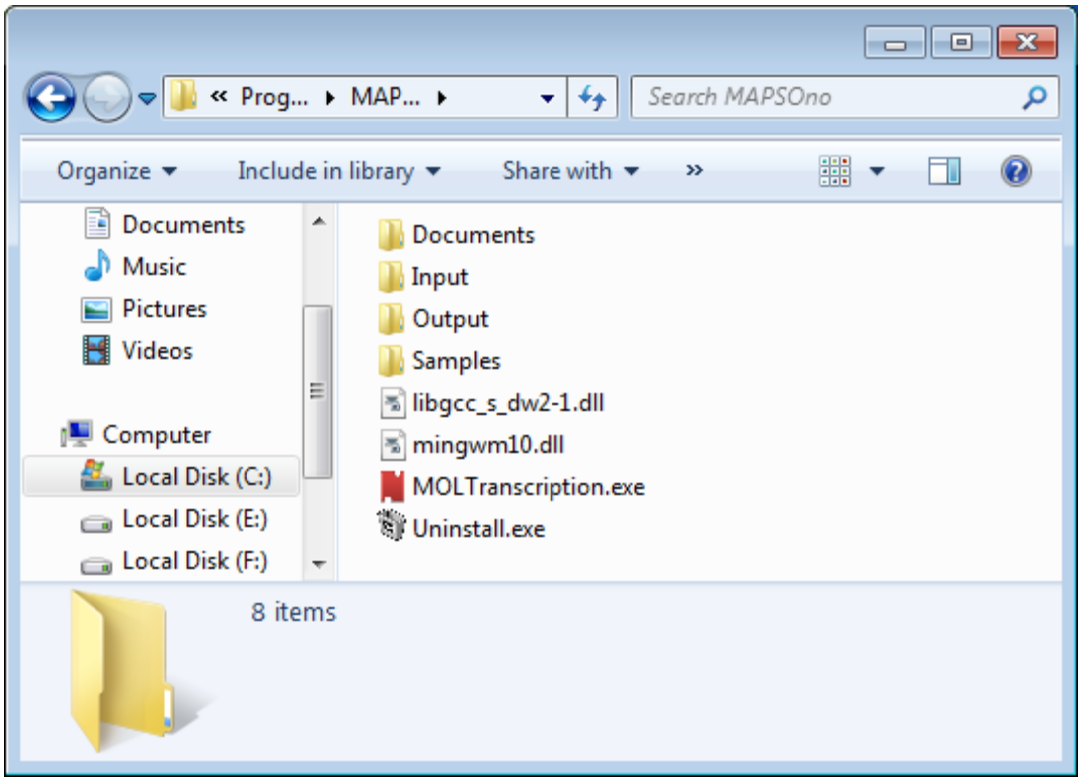


Figure 1

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

The program will create a subfolder "MAPSOno" in the folder "Program Files" and additional folders with names (Input, Output, Samples, Manual) as well as the basic program files and the program icon on the desktop.

The program does not add any entries to the Start menu "Start Menu" or entries to the system registry, so it can be removed by deleting files only, the software comes with an unistaller tool for this purpose described in the section how to remove the program from your system.

Uninstalling the program

To remove the program, go to the program folder then run the uninstalling utility (Uninstall.exe) by double-clicking, you can do so through the following steps:

1. Right click on the program icon on the desktop to display the properties popup menu and choose (Properties) at the bottom of the list, you'll see the screen shown in Figure 2 below.
2. Click on the button (Find Target ... or "Open File Location") to go to the program folder, the contents of the folder may be similar to what is shown in Figure 3.
3. Run the uninstaller program by double-clicking on the icon (Uninstall.exe).
4. You will need to remove the program's icon on the desktop and folder "MAPSOno" manually. Please note that "MAPSOno" folder may contain other programs made by Kalmasoft so, unless you are sure, leave it untouched.

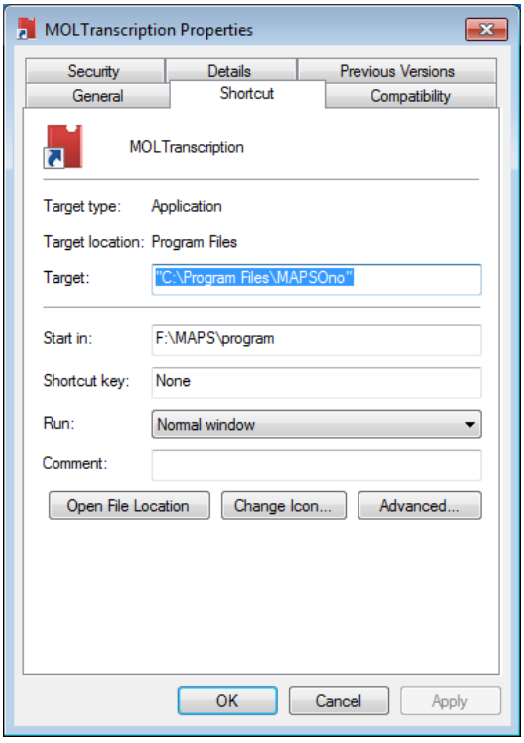


Figure 2

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

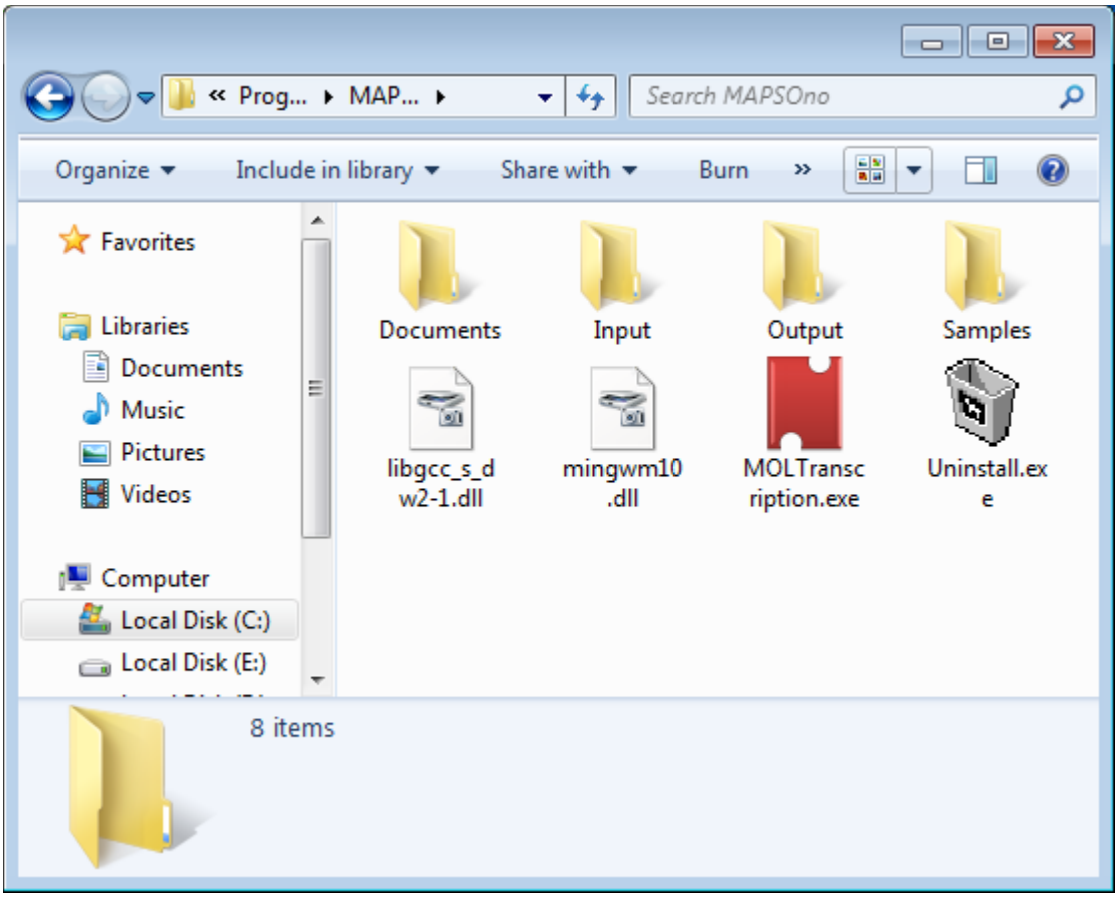


Figure 3

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Program folder contents description

1. "**Input**": contains Arabic personal names text files you load to the program, this folder will be searched for any text files that may contain personal names to load to the program, unless you specify another location for input, the program opens this folder by default.
2. "**Output**": contains transcribed names output files, this folder will be targeted by the program for output, unless you specify another location for output, the program will use this folder by default.
3. "**Samples**": sample input files can be found here containing some Arabic personal names for the purpose of testing the program.
4. "**Documents**": contains program documentation, you'll find all the important documents and notes in this folder, including the full text copy of this manual in Adobe PDF format.

The rest of the files include the executable file of the program, uninstaller tool in addition to other files and libraries necessary to run the program.

When you run the MAPSOno® Lite (Transcription) for the first time the Input Interface (Figure 4), a panel through which you can enter Arabic personal names, will appear, enter few names and then click on the (Transcribe) button at the bottom of the interface, this will immediately start the transcription process then the program will automatically switch to the View Interface (Figure 5) to show the result of transcription.

Transcribed names will be displayed according to the target language script together with some other important information including identification numbers, the orthographic status of each name, messages manifesting the errors if any, and hints to give an idea about the type of those errors and how to deal with them.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

A quick start

The following steps summarize how you can use the program, we recommend you start from here, then read about the details of program interfaces in the following paragraphs, and then move to the explanation starting from the paragraph (Running the program) on page (35).

1. In the Input Interface, enter or paste a list of Arabic personal names in the input pane, make sure that you press (Enter) at the end of each name when typing, the input pane is positioned at far right side on the Input Interface.
2. Click on the (**Transcribe**) button when you done, the program will automatically switch to the View Interface and you will see the results of the transcription on the result display area. By default, the input names will be transcribed to English but you can select from other 25 international languages supported by the program.
3. You can return to the input pane to make necessary editing or add some new personal names.
4. The program comes preset with some default values but you can change them as necessary as soon as you are fully aware of the nature and functional aspects of the program.
5. Select another target language from the master system selector menu, the program will again process the names and display the results according to the new language you selected.
6. In the first column you'll see icons indicating the status of each name, hover over the icons to view a tip of what each icon stands for, the hint in the last column will tell the reason of why the icon is there. If the transcription process is successfully done, the icon will show a green filled bullet otherwise you will see different icons of different colors and you'll find a corresponding clarification about the status of the name or the type of the error a name may have.

You can then output the result to the printer or save it to a file for later use, these two processes (printing and saving) are so conventional thus no further explanations is given in this manual.

The following pages will explain all the steps above in detail starting with the definition of the interfaces of the program.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Program interfaces

MAPSOno® Lite (Transcription) has five interfaces, Input Interface to enter Arabic personal names, View Interface to display transcription results, Settings Interface to control the program behavior, Fine Tuning Interface, and a help interface for documentation on how to use the program and other important information.

You can navigate between the interfaces by clicking the appropriate interface tag at the top; some important parts will always be present and displayed with all interfaces, these include program's master function buttons, Quick Options pane, and the status pane.

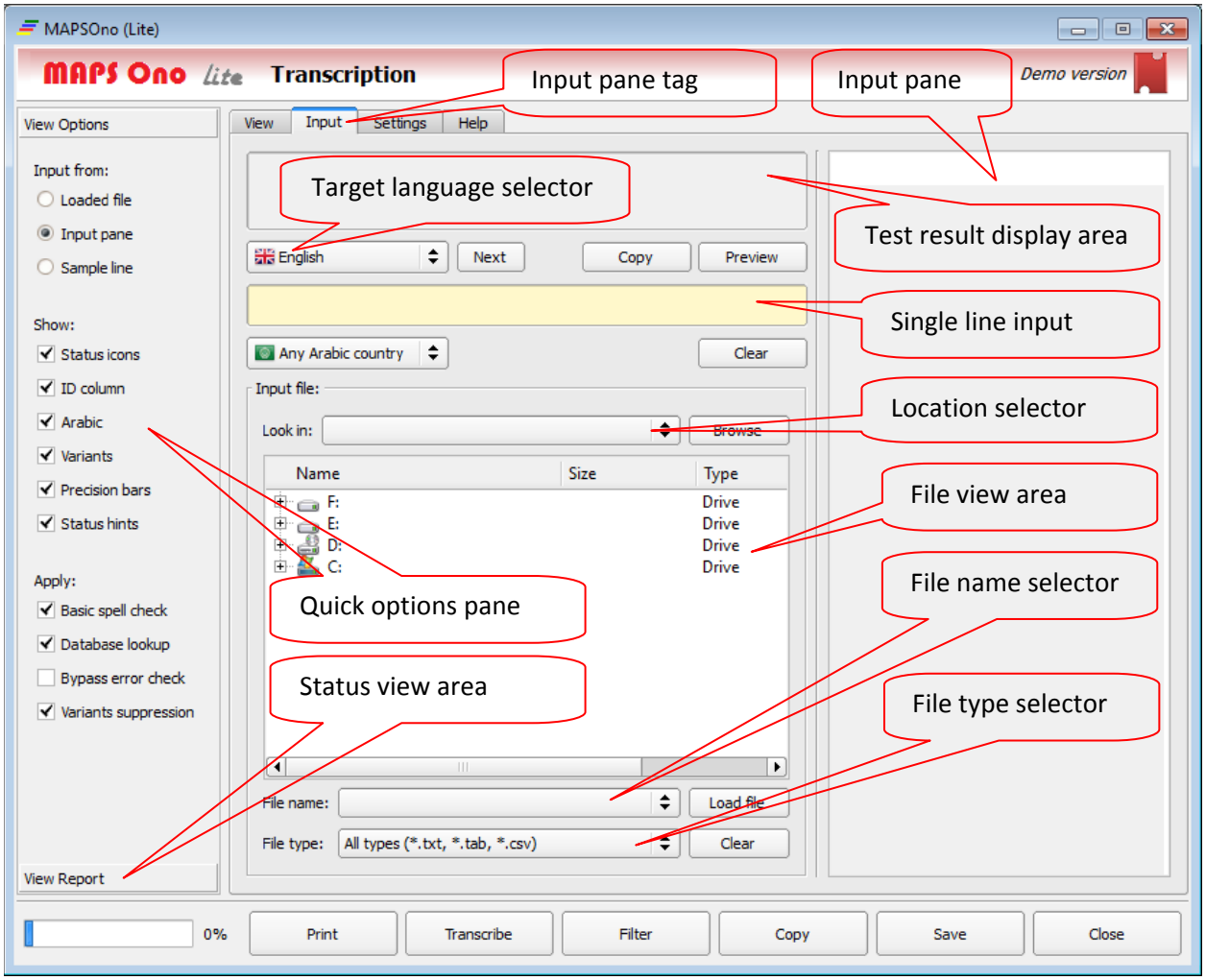


Figure 4

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Master function buttons

- (**Print**) prints the contents of the result display area, your operating system may provide other print options such as printing to a file or to Postscript for example.
- (**Transcribe**) to start the transcription process of the input Arabic personal names.
- (**Filter**) to search for a particular name or to sort the names according to specific criteria.
- (**Copy**) copies the entire contents of the result display area to the clipboard, so you can then paste it to other applications for further formatting as necessary.
- (**Save**) to save the results in a particular format text files or other types of files.
- (**Close**) closes the program.

Quick options pane

Quick Options pane contains options specific to the input and View Interfaces, these are classified to the following three groups:

Determining the source of personal names:

Click on the appropriate radio button to switch between the three sources of personal names (load from file, from the input panel, or the single line input).

Customizing the result display area:

You can customize the display area by adding or hiding columns, click on a check box as appropriate. Columns that can be hidden are listed in the pane.

Controlling the behavior of the program:

You can rapidly intervene in the way the program process Arabic personal names by placing a check in the appropriate box where you can impose the spell checker, use the integrated database, or disable the variant generator; you may also bypass a limited number of potential errors in order to continue with the transcription process.

Transcription progress bar

This is a dynamic progress indicator bar, it shows the percentage progress of the transcription process.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Input interface

When you run the program for the first time (and every time) it will automatically start in the Input Interface as shown in Figure 4 above.

The lower part of the interface contains the master function buttons to perform the basic operations such as copying, saving, and printing, these buttons will remain displayed with the other interfaces. To the far left side are the Quick Options pane and the status display pane/area, they will always be displayed too.

Buttons

- (**Copy**) copies the result of transcription for a single name from the test result display area.
- (**Preview**) displays the result of transcription in the test result display area.
- (**Next**) displays next transcription variant of a single name.
- (**Clear**) clears the single line input, the lower "**Clear**" button will clear the list of names from the input pane as well as the memory from the loaded input file. Clearing the memory of the computer will have no indication except for a message shown in the status display area. This procedure is necessary if the response of the program become slow or you do not need the loaded names anymore.
- (**Load file**) to load a file of personal names.

Menus

- Target language selector: use this menu to set the transcription language for testing a single line name.
- File location selector (Look in): use this menu to locate the file folders.
- File name selector (File name) use this list to select one of the previously loaded files.
- File type selector (File type) use this menu to determine the type of file; generally speaking, files that can be loaded by MAPSOno® Lite (Transcription) are text files, they may differ in the format and encoding, which are to be determined in advance, however, adjusting the program to accept specific file format and encoding is described in the paragraph (Input file format and encoding pane) on page (25).

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Input pane

Use the input pane for an average of 1500 names, names can be input in several ways:

- Type directly using the keyboard.
- Paste a copy of names from another application.
- Drag and drop contents from another application window or text editor.

To clear all the names click on the (**Clear**) button at the bottom, you can also highlight or select all the contents by pressing on (Ctrl + A) and then hit the (Delete) key from the computer keyboard.

Names can be separated from each other in several ways:

- Each name is in a separate line, the delimiter in this case is (\n) mark i.e. the new line, this is the default delimiter.
- Adding any of the following characters (- , _ ; / \) or a space or tab, in this case you must specify the type of the delimiter from the setting interface, see paragraph (Setting the names delimiter) on page (36) for more details.

Single line input

The single line input is for testing purposes prior to Transcription a large number of names, use this input to enter one line of text representing the name in the following way:

1. Type the name or paste a copy from another source, you may also drag the name from the input pane to the right.
2. Choose the Transcription language using the transcription language selector.
3. Click on the (**Preview**) button to display the result of transcription, the result will be displayed on the area at the top; press the same button every time you change the transcription language to get a different result.
4. The (**Next**) button will be enabled if the name has orthographic variants and disabled when the navigated list comes to an end, click the (**Preview**) button again to return to the first variant.
5. Click on the (**Clear**) button to clear the input name.
6. You can copy the result of Transcription shown on the test result display area by pressing the (**Copy**) button, paste to any other application, and make sure that you have the correct font to display all the letters and symbols.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Load file pane

You can load a file of geographical names using this pane, the source file will not changed until you load another file or remove the file by clicking on the (**Clear**) button located in the pane.

View interface

The view interface is the program's main interface, shown in (Figure 5) on the next page, most of the transcription operations are controlled from this interface; components of this interface are described below.

Result display area

Centered in the interface, the result display area occupies most of the space to display all the results of operations done by the program; all results are presented in a form of an adjustable table with contents can be sorted alphabetically.

- Transcription results can be copied directly from this interface by selecting the item or line and then pressing (Ctrl + C).
- To select the entire row (name + the results of different transcription languages), click on the row header to the left or at the row number.
- To select the entire column (the results of a single transcription language) click on the column header.
- To select two or more columns (all the names + the results of specific transcription languages) drag the mouse on the headers of the selected columns.
- To sort the contents of any column, press the left mouse button on the column header, this will alphabetically sort the contents of the entire table; you can reverse the sort order by clicking on the column header again.
- To adjust the column width drag the column border at the header to the left or right.
- To adjust the row height drag the row border at the head up or down.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

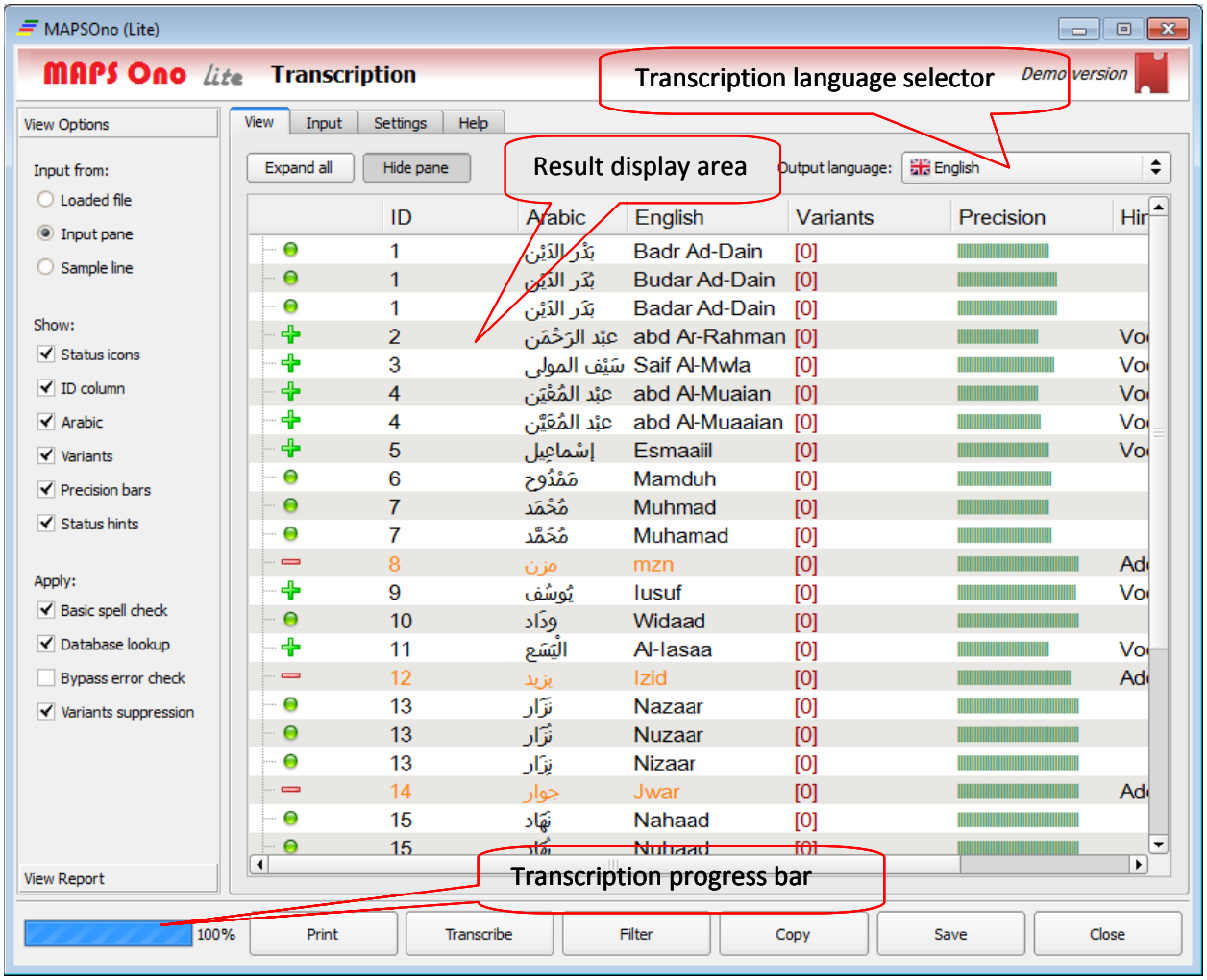


Figure 5

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Master language selector menu

A drop-down menu at the top right corner of the View Interface, serves as the master control menu of the different transcription schemes. You need not to click the (Transcribe) button every time since this selector will trigger all necessary actions in a single step.

Do not use this menu to rapidly switch between transcription languages if the number of names to be transcribed is very large (hundreds of thousands), do your tests using a limited number of names first so you will not have to wait a long time before you get the results.

The items of this menu include all the 25 transcription languages supported by the program in addition to the following options (may be optional):

Option	Details
Default Language	Names will be transcribed using the default language (Pro version only)
Currently selected languages	Names will be transcribed using only selected languages
All supported languages	Names will be transcribed using all supported languages (Pro version only)
None, check mode	No Transcription , select for spell check

Status display area

Status display area is where reports and comments related to the active interface appear, messages concerning the status of the program and the kinds of ongoing operations will also be displayed here; some hints will also be shown, the details are as follows:

- Statement of the input source of the personal names.
- A brief definition about the function of the active interface.
- The status of the program during the processing of personal names.
- A mini statistical report about the input names, the operation conducted, number of errors detected, and alerts issued; the ideal situation is to get 100% of the total names transcribed without any errors or alerts, but this does not reflect the perfection of the accuracy of transcription because the program does not perform according to phonetic constrains, it only does a lexical scanning followed by the basic phonemic transcription. The only way to tell whether the transcription is good is through the Precision indicator bars, please refer to page (45).

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Error messages and status hints

The program displays enough information about the status of each name after the Transcription process; hints may either reflect the status of the Transcription process or inform the user that some changes has been made to the name. Most of the errors can be overcome by simply ignoring the error, but some names cannot undergo the process of Transcription because of errors hard to overcome or correct. You can refer to (Appendix A. Error messages and status hints) on page (57) where a complete list of types of errors are manifested.

Types of errors that can be identified include:

1. Inconsistent vocalization: for example the position or type of the diacritics.
2. Spelling: location and type (Arabic or non-Arabic), e.g. improper placement of letters.
3. Mojibake: noisy names that contain non-Arab characters, digits or other symbols.

> CAUTION: The hint always describes the first error (حذيفُة عيسُى) for example has two errors, "Ḍamma on the letter Nuun and letter Seen" only the first error will be manifested (the Professional Edition shows all errors and their respective positions).

- MAPSOno® Lite does not fully suggest error correction or allow doing correction on place from the view interface directly (the professional edition provides all of that), retype the correct names to get rid of the errors messages.
- If the table cells do not show the full text of the hint you can move the border around the edges a bit to make more space for the full text, drag the right border to enlarge the cell width.
- You can check the appropriate box in the quick options pane to hide the status icons, the whole column will rapidly disappear, remove the check and the icons will appear once again.
- Identification numbers could be hidden from the quick options pane in the same manner as above, check the box [*ID column*] the second column will disappear, this is a useful feature serves as a reference for later proofreading.
- Hide the hints column from the Quick Options pane by checking the box [*Status hints*].
- The whole Quick Options pane can be hidden by clicking on the (**Hide pane**) toggle button, this is the best solution to get a wide view area of the results.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Settings interface

The settings interface is where MAPSOno® Lite (Transcription) overall behavior can be controlled, resources of the program and handling strategies of Arabic personal names can be set from here; we will first describe the main action buttons and then describe each part separately in detail.

Main setting buttons

- (**OK**) accept and save the new settings.
- (**Restore**) restores the default settings.
- (**Fine tuning**), a toggle button switch to Fine Tuning Interface. Click again to switch back to this basic settings interface.

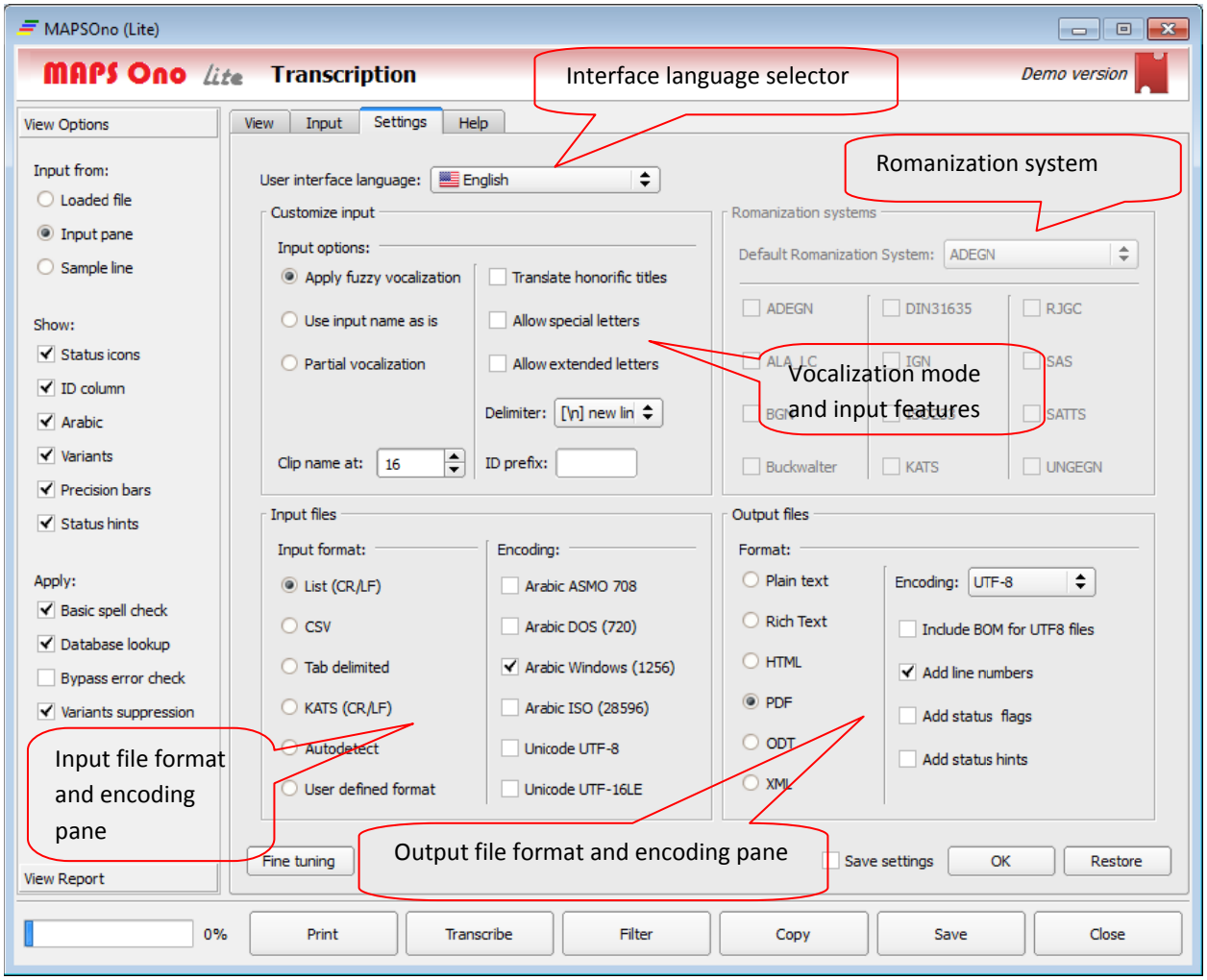


Figure 6

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

To understand how this interface works split vertically into two parts, the right part is for output, the left is for input; further divide each part into upper and lower panes, the upper pane is for the display, the lower is for files. So all controls regarding files and print can be found at the bottom of the interface and everything related to the name input and display on the top.

We will begin describing the interface panes clockwise starting from the top left corner.

Layout direction and widgets placement

This manual is based on the English interface, if you would like to see all panels placed from right to left please switch to Arabic/Hebrew language and you will have the right orientation.

User interface language selector

You can select from the sixteen interface languages supported by the program at any time, the interface will change dynamically as soon as you select a language; languages are written in their native scripts so to facilitate the search for your preferred language in case you missed your native language by mistake.

Source language pane

Arabic language basic setting is done from here, you can set some essential features like mode of vocalization, name format and few other related features which are discussed below:

Vocalization mode selection pane

This pane has two parts; the left pane contains three options and a spinner, described as follows:

- Apply the Fuzzy vocalization mode (default) to bare or unvocalized names; Fuzzy Vocalization depends on morphological templates and statistical criteria as well as of some basic spelling rules.
- Use the name as is, that is, without any help from the program, use this option only if you are sure the names are fully vocalized personal names.
- Partial vocalization, this is the default mode and the program will help you by automatically completing the necessary vocalization if you just add the minimal disambiguating diacritics.
- Determine the length of the name, the maximum number of characters of each name shall not exceed 20, the minimum is 3, spaces and diacritics are counted.

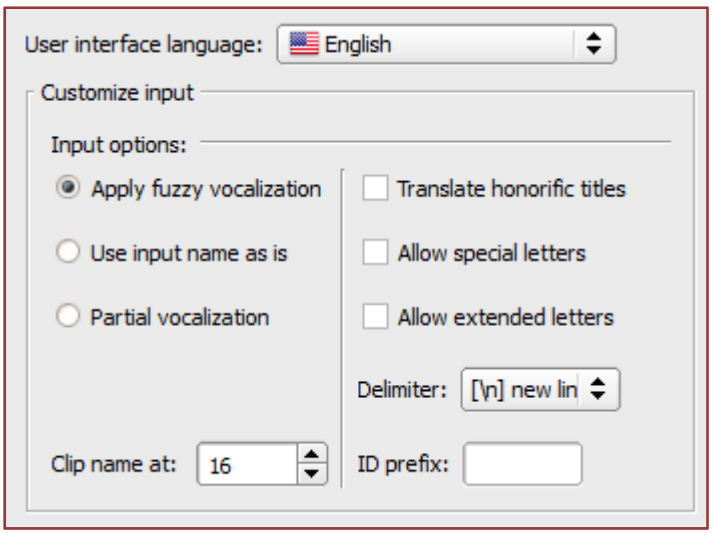


Figure 7

The above vocalization modes are described in details in paragraph (Modes of vocalization) on page (38).

The right pane contains three options, a menu, and an input control described as follows:

- Translate honorific titles that may precede a given personal name.
- Allow the insertion of special characters in names, such as letters (ڤ, ڃ, ڤ).
- Allow the insertion of Arabic extended letters used in other languages, such as Urdu and Farsi, e.g. (ڭ, ڭ).
- A delimiter selection menu.
- An input to set the identification number prefix of personal names.

See the details of how to use this pane in paragraph Preparing the names for input on page (36). You can edit any name in place before the transcription process since the input pane serves as a text editor and supports all basic functions such as cut and paste, drag and drop, and redo functions by using the Microsoft Windows® usual shortcut keys.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Input file format and encoding pane

Only one file can be set using this pane each time you load names from a file (the professional edition allow for setting multiple files in batch mode)

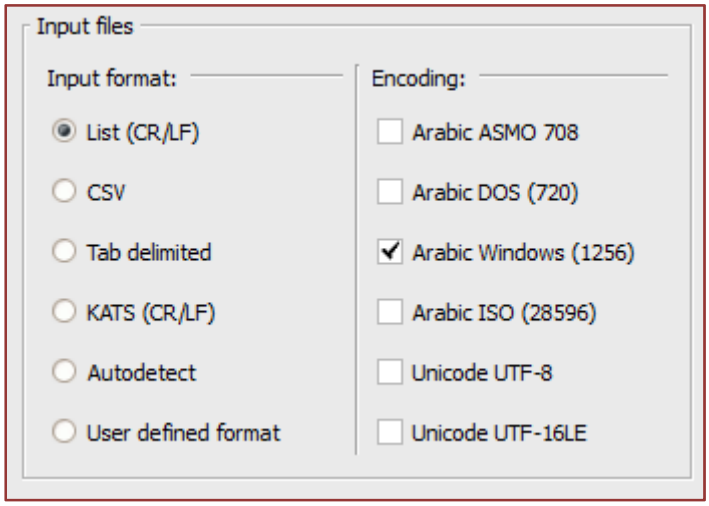


Figure 8

Supported formats are shown below:

Encoding	Details
Windows 1256	Default encoding for Arabic in Windows operating system
UTF-8	Unicode encoding
UTF-16LE	Unicode encoding

Format	Details
List (CR/LF)	Names formatted in a single column and separated by newline
CSV	names separated by a comma
Tab delimited	Names separated with the Tab character
KATS (CR/LF)	Format based on Kalmasoft KATS
Autodetect	Format is automatically determined by the program
User defined format	Format specified by the user

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Output file format and encoding pane

Output file setting should be prepared before the saving the results of Transcription because the program uses some symbols and characters that may not be supported by all types of user viewing client software. MAPSOno® Lite supports Unicode fully in both input and output so please adopt this encoding to ensure that all details are preserved.

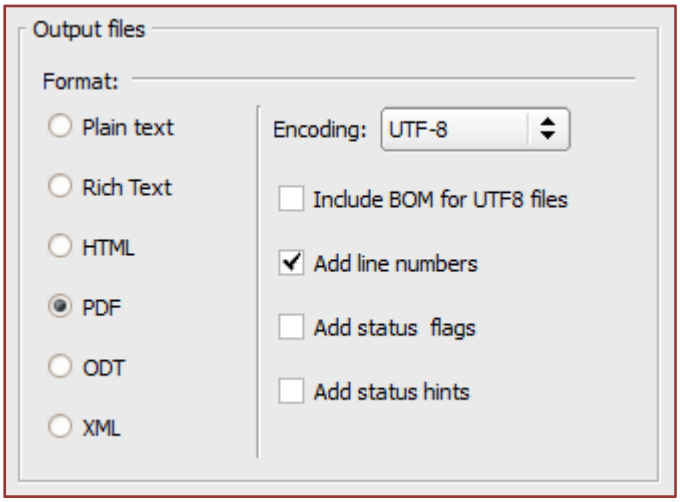


Figure 9

The file format depends on how the result will be used, for viewing and printing purposes you can use any of the formatting types (Rich Text, HTML, PDF, ODT, SQL), but if the results will be subject to subsequent processing then you must use any of these formatting (Plain text, XML) instead.

File encoding is subject to the type of application that will be used later to display the results or make subsequent processing, in most cases choose (UTF-8) with the box (Include BOM for UTF8 files) checked. You can also see a brief description of each encoding system in the built-in help guide with included in the program.

Encoding	Uses	Example of applications
UTF-8	for most multilingual applications	
UTF-16LE	use with specific applications	Microsoft Access-2007 Excel-2007,
UTF-16BE	use with most applications	Apple Mac OS X
Windows-1256	use with all Microsoft applications	

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

The remaining options are self-evident and include the addition of line numbers to the file, status flags, as well as addition of the status hints necessary to review all errors or make a deep examination by the user.

The program provides output files in four different forms of encoding, encodings can be controlled from the dropdown menu provided for different uses as summarized in the table below.

Format	Full name	Details
Plain text		Processing purposes
Rich text		Display and print purposes
HTML	Hyper Text Markup Language	Web pages
PDF	Portable Document Format	Display and print purposes
ODF	Open Document Format	Display and print purposes
XML	Extended Markup Language	Processing purposes

Romanization system selection pane

Select the preferred Transcription languages by checking the appropriate boxes, you can see the characteristics of each system in the table below. You can also see a brief description on each system in the built-in help guide that comes with the program package.

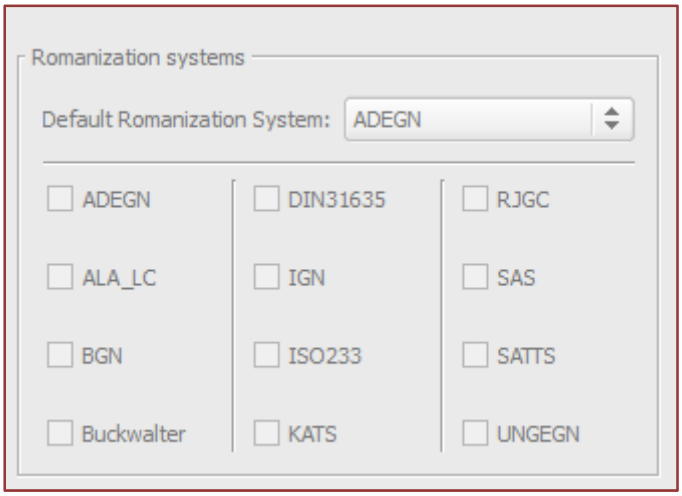


Figure 10

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

System	Transcription	Full name
ADEGN	Transcription	Arabic Conference on Geographical Names
ALA-LC	Transcription	Association of American Libraries -Library of Congress
BGN	Transcription	Board of Geographical Names/ Permanent Committee on Geographical Names
Buckwalter	Transliteration	Tim Buckwalter Arabic Transliteration
DIN31635	Transcription	Deutsches Institut für Normung
IGN	Transcription	Institut géographique national
ISO233	Transcription	International Standards Organization
KATS	Transliteration	Kalmasoft Arabic Transliteration System http://www.kalmasoft/devtool.htm
RJGC	Transcription	Royal Jordanian Geographic Center
SAS	Transcription	Spanish Arabists School
SATTS	Transliteration	Standard Arabic Technical Transliteration System
UNGEGN	Transcription	United Nations Group of Experts on Geographical Names

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Notes:

- If your operating system does not support Arabic language, it is preferable to use the (KATS) format for input.
- If the operating system you have is different from Windows, you can run the program on a Windows platform, save the files using (UTF-8) encoding and in any suitable format, the UTF-8 is a global encoding for files and most platforms support it.
- Input files may not be encoded in (Windows 1256), in this case, make sure the file encoding is supported by MAPSOno® Lite (Transcription), set as appropriate from the settings interface; if no way to recognize the encoding, set the program to (Autodetect), the program takes time to detect the encoding before displaying a sample of names on the input pane.
- Make sure that the format specified in the program is the same as the input file format, there is no way in the program to check it out, there are some text editors you can use to manipulate your data format.
- If you do not have a professional text editor save the list of names to Microsoft Excel worksheet then save as "Text Tab delimited" or "CSV comma delimited"; these two formats are supported by the program.
- In most cases, you can simply have your list arranged in a column in a Microsoft Excel worksheet or as a column of a table in Word or even a simple list using Microsoft WordPad; copy and paste the names to the input panel.
- If you use the Microsoft WordPad, make sure you do not save the file in "RTF" format, save it in "Text Document" format, MAPSOno® Lite cannot deal with Arabic geographical names stored in "RTF" formatted files.
- If you use the Microsoft Notepad, make sure to save your files in "Text Documents" and in "Unicode" encoding, Notepad will store the file in (UTF-16LE) which is supported by MAPSOno® Lite, you can then set the input as appropriate.
- You cannot re-open a file containing Arabic geographical names and Transcribed names, only the professional edition can open files with mixed contents.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

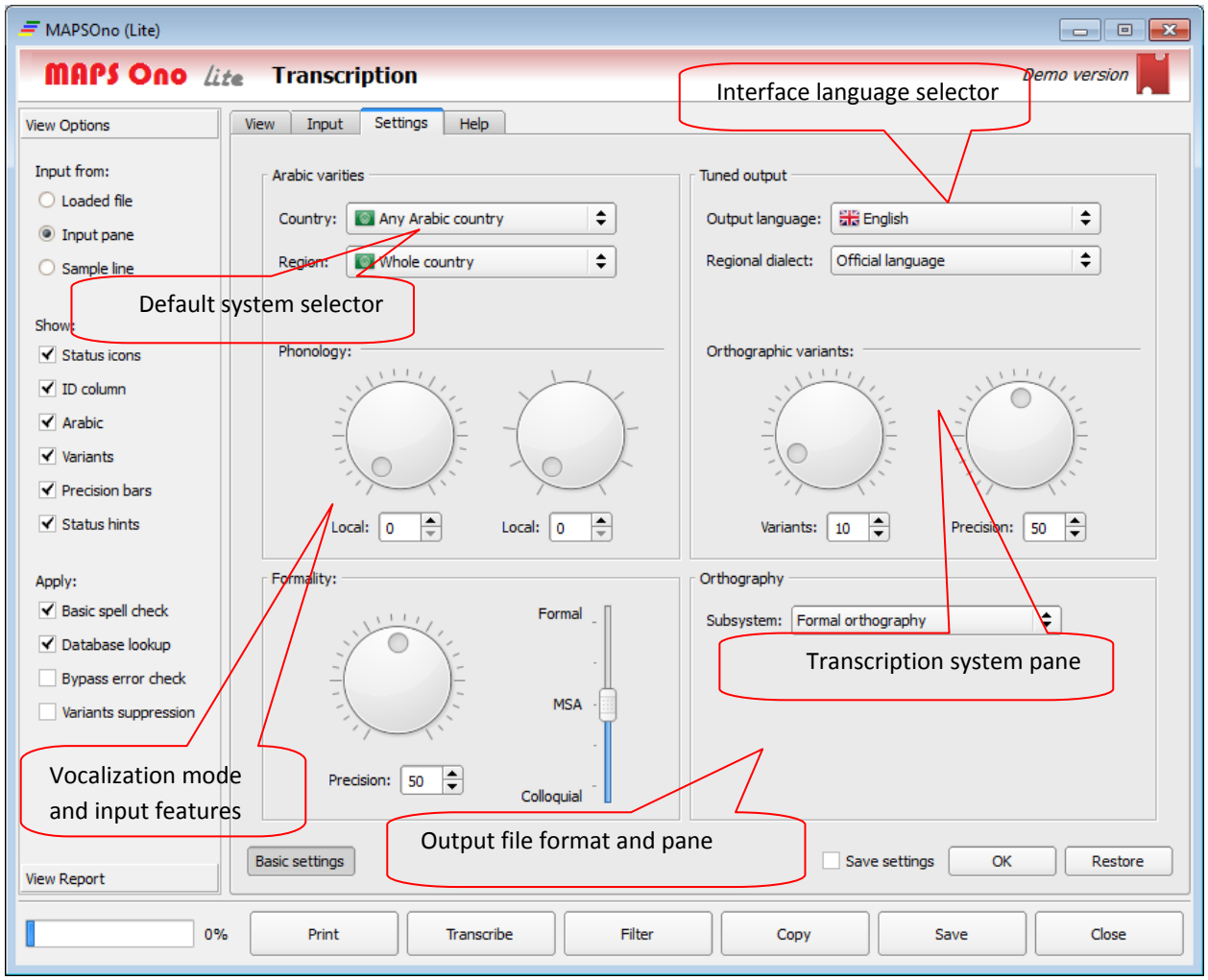


Figure 11

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Fine Tuning Interface

Both Setting Interface and Fine Tuning Interface can be accessed by clicking on the (**Fine tuning/Basic settings**) toggle button; the Fine Tuning Interface is where you can adjust the minimal details and direct some specific aspects that have subtle effects on the transcription process.

Arabic language settings

Arabic language has varieties, you can have the program considering these varieties, this will affect the way some personal name are transcribed; you can even specify the geographic region within the country, please note that the list of countries includes members not in the Middle East e.g. Pakistan, those are countries using Arabic Script Based Languages (ASBL); unless you know what you are doing please leave this part in the default setting i.e. (Any Arabic country).

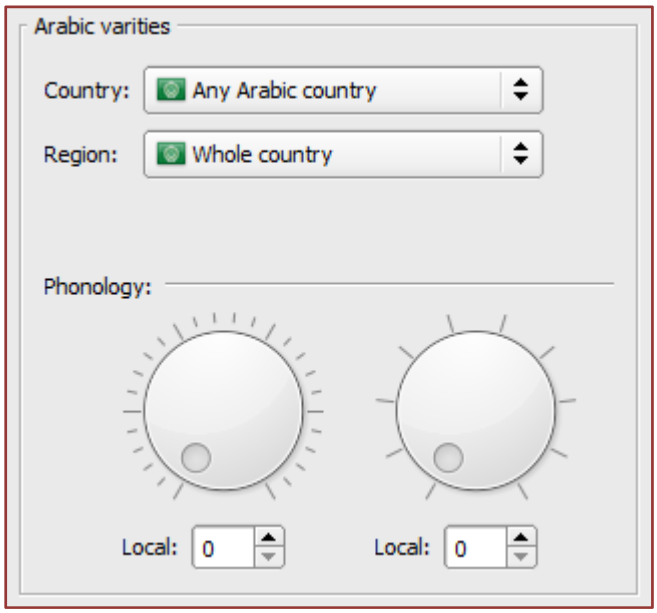


Figure 12

Arabic language formality level

By default, the program transcribes names based on the Modern Standard Arabic "MSA" but you can select either formal Arabic "*Fu'sha*" or colloquial Arabic "*'Amiah*"; this is shown in (Figure 13) below. The effect is that transcribed names will follow strict Arabic vocalization conventions e.g. *Tanween* and final gemination as in (Ali > Aliy) in the case of formal Arabic, and will follow a very custom regional pronunciation rules if you select "Colloquial".

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

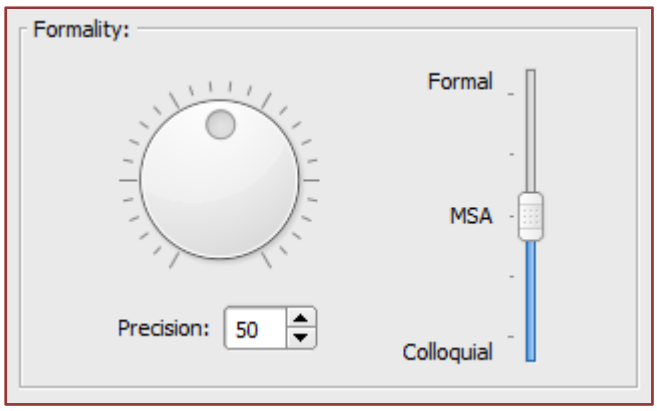


Figure 13

Target language varieties and regional settings

Some languages have regional dialects, the program will let you deal with a limited set of those dialects by setting the [*Regional dialect*] menu from the Fine Tuning interface, in case you have no idea how to deal with this menu please leave the default value i.e. [*Official language*] intact.

Target language regional settings may have no transcription effect but on few input names, it only clear in cases where the official and the regional dialect sound systems are different enough e.g. English and Welsh, Spanish and Basque.

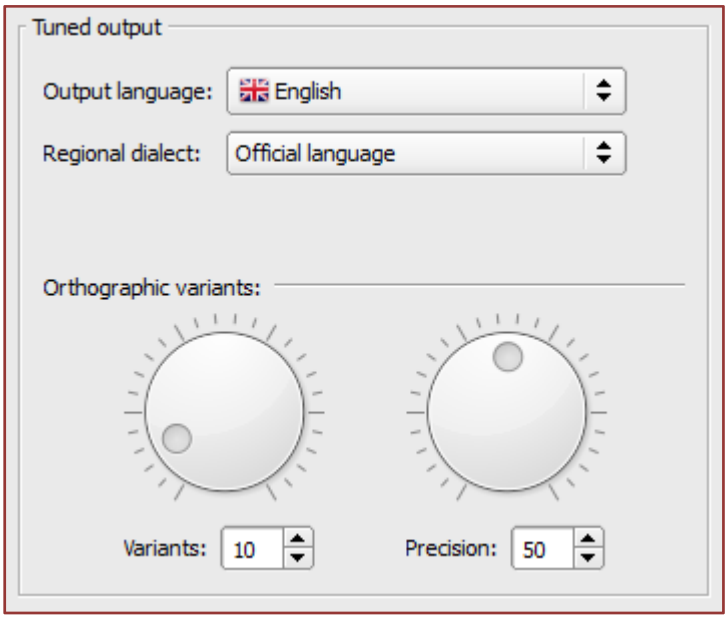


Figure 14

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Orthographic variants settings

This part is discussed in detail in chapter (Orthographic variants); the panel can be seen in (Figure 14); setting the precision dial to 50% and the variant number dial to 10 will limit the variant *displayed* to only those *first* ten candidates passed the 50% level; all variants will be displayed in case the total number *generated* is less than the set number provided that they all above or equal 50% precision level.

Target language writing subsystem

Showing the results in specific writing system is also possible, for the Japanese language the default is *Katakana* but you may select *Hiragana*; *BoPoMoFo* is the default for Chinese but *Pinyin* is an option too.

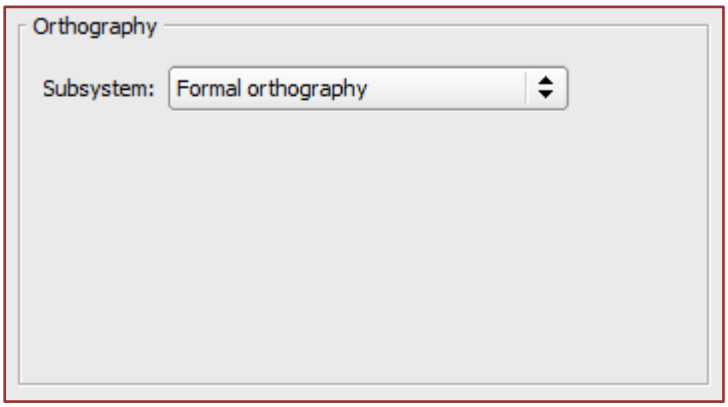


Figure 15

Program default settings

MAPSOno® Lite (Transcription) comes preset with default settings but you may change these settings at any time during or before the transcription session as needed, you can also save your settings for future use or restore the initial situation when you installed the program for the first time, the following table shows the status of all these settings:

Interface		Value / setting	Details
Input	Single line transcription	English	
	File preview	All drives	
	Input file location	MAPSOno/Input	Set as required
	Output file location	MAPSOno/Output	Set as required
View			
	Master system selector	Current transcription	
	Default transcription	English	
Setting			
	Interface language	English	
	Default Romanization system	ADEGN "Arabic Conference on Geographical Names"	
	Other transcription language	Not checked	Check as required
	Mode of vocalization	Fuzzy vocalization	
	Length of name	10 characters	Maximum is 25 characters
	Default delimiter	\n	New line
	Input pane format	List (CR/LF)	Name per line
	ID prefix	Nothing	
	Allow digits with names	Not checked	Not allowed
	Allow special letters	Not checked	Not allowed
	Allow extended letters	Not checked	Not allowed
	Maximum displayed variants	10	
	Maximum variant precision	50%	
	Input file encoding	Arabic Windows®-1256	Windows® standard encoding
	Output file encoding	UTF-8	Unicode
	Input file format	List (CR/LF)	
	Output file format	PDF	
	Include BOM for UTF box	Checked	
	Add line numbers box	Checked	Line numbers will be added
	Add status flags box	Checked	Status flags will be added
	Add status hints box	Not checked	Hints and tips will be added
Program			
	Quick Options pane	Shown	
	Names input source	Input pane	
	Hide check boxes	All not checked	
	Basic spell check	Checked	
	Bypass error check	Not checked	
	Use database	Checked	

Saving user custom settings

The program provides the ability to save the user custom settings, click on (**OK**) button and your current settings will be saved for future use, these will be persistent and affect the program's behavior each time you run it unless you modify or apply new settings.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Running the program

Setting the program for input

MAPSOno® Lite (Transcription) does not require any special fonts, but you have to make sure that a single font is available and already installed in your system (Arial Unicode MS), this font usually comes with Microsoft Office®, it an Unicode font, meaning that it contains all characters and symbols used by many different languages supported by the program. To make sure your system has the required font type/paste the word (رَشيدة) in the single line input pane and select the transcription language (let it be Japanese for the sake of testing), you should see the word (ラシーダ) the correct transcription for "/raʃi:da/" on the test result display area as shown in the example below.

OK, your system is ready	ラシーダ	رَشيدة /raʃi:da/

Arabic personal names can be input from three different sources, single line input for testing, input pane for moderate lists of names using the keyboard or copy and paste techniques, and by loading files of large numbers of names. When you are in the View Interface, you can quickly switch between these sources as necessary, in fact you can actually put names for processing in all these three sources and process each one after the other (professional edition allows the insertion of multiple files and process all at once).

The master language selector menu in the View Interface determines the current transcription language, it provide necessary functions to transcribe the personal names according to user settings as well as all Romanization systems supported by the program in one single step.

It is not necessary to arrange the names in a vertical list, you can enter names separated by any of the delimiters described in paragraph (Setting the names delimiter) on page (36) prior to the transcription process, the program will portray the names in a list format for you to confirm the input, if you use the default delimiter "newline" you will not be asked for confirmation. You can edit any name in place before the transcription process since the input pane serves as a text editor and supports all basic functions such as cut and paste, drag and drop, and redo functions by using the Microsoft Windows® usual shortcut keys.

The master language selector works simply by selecting the item so you do not actually need to click the (**Transcribe**) button every time to start the transcription process, if the list of names is too long (few hundreds of thousands) you will have to wait a short time before the results show up; we recommend you make necessary setting in advance in this case, then click the (**Transcribe**) button to initiate the process. This master selector is designed for your convenience and to facilitate rapid action, you can also use it to bypass the settings you have made in the settings interface except for the possibility of changing the default Romanization system, that must be carried out only from the Setting Interface.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Preparing the names for input

MAPSOno® Lite (Transcription) is very sensitive to the input data, it carries out several tests to before the transcription process begins, these tests generally include spell checking for properties such as length of the name, type of characters and the type of vocalization, the program also does some other covert arrangements include the removal of unnecessary characters such as white spaces and the Arabic lengthening characters "Ṭaṭweel".

The program provides many features for the user to apply on the personal names before transcription, all can be applied from the settings interface, the available features are as follows:

Setting the maximum length of the name

User can set a precondition for the length of the input name, MAPSOno® Lite (Transcription) default is (10) characters, including diacritical marks and spaces, the user can set this to any number between 3 and 25; this will add an indirect means for correcting the names of excessive length.

No special icon will be assigned to clipped names because other icons are still needed to reflect the status of the remaining part of the name, a scissors (✂) flag is incorporated directly with the Arabic name to indicate that it has been cut short, this flag will not be printed or saved to file, instead, a tag word (clipped) will appear in the hints column, the example below shows a name clipped after a maximum length of 22 characters.

Alaa-Uddiin Muhammad M	عَلاء الدين مُحَمَّد م✂

Setting the names delimiter

When using the input pane to enter the personal names you must specify the delimiter that separates the names; the program is set to use a common delimiter (end of the line) that simply means you press the (Enter) key after each personal name you type so you will have the names arranged in a vertical list (column), but you can change the delimiter to suit input data arranged in different configurations, for example the following arrangement is acceptable:

محمود، صلاح، المازني، بابكر، مصطفى، نهاد، أيوب، موزة، حصة، البتول

The above example shows the use of the comma (,) as a delimiter, you can also enter the names separated by any of the following characters: (";", "tab", "_", "-", "\"), and the space. If you chose the space make sure that you type the parts of a compound name in between parenthesis like this (عبد الرحمن), or connect them with a hyphen like this "عبد-الرحمن" so the program will not parse each part separately.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Setting the ID prefix

Use the ID prefix to decorate the identification numbers of the personal names, for example, if you want to transcribe names from different geographical areas for subsequent processing the ID prefix will be useful in sorting, to use this feature type a word of maximum four characters in the appropriate place, for example "IRAQ" to get the ID numbers prefixed in the following format (IRAQ005417) or "KSA-" to get (KSA-005417), the numerical field width is six digits enough for one million names, the example below shows the use of the ID prefix "KSA-".

KSA-005417	الغامدي	Al-Ghamidi

Honorific titles translation

MAPSOno® Lite does not translate titles by default but you may activate this feature from the Settings Interfaces, see paragraph Modes of vocalization on page (38), we generally do not recommend using this feature, the example below shows how this feature works.

Prof. Abdallah	البروفيسور عبد الله

Allowing the insertion of special letters

Arabic letters such as "چ" (*Cheh*) "ڤ" (*Qaf* with three dots above) are sometimes used with names of non-Arab origins, MAPSOno® renders these characters to closest Arabic letter from the script if not supported by the transcription language specified.

at-Tājī	التاچي
Akjūjit	اكچوچت
Sīdī Vāl	سيدي ڤال

Allowing extended Arabic letters

MAPSOno® Lite allow using letters used by Arabic script based languages (ASBL) such as Farsi and Urdu which use extended Arabic letters, some writing scripts in Morocco uses them as well, these letters are slightly different (ڭ, ژ, گ), we do not recommend that you allow insertion of such letters unless you are certain how they can be used.

Jangīz	چنگيز

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Modes of vocalization

MAPSOno® Lite (Transcription) provides three modes of vocalization, these can be set from the Settings Interface, a summary of these modes shown here with examples, a detailed description can be found in the full text manual.

Fuzzy vocalization

Fuzzy vocalization is used to vocalize bare or unvocalized Arabic names, it adds diacritics to the names using heuristics and some statistical profiles, the process depends on the type of letters in general; as a consequence, fuzzy vocalization generates orthographic variants, this feature is integrated in MAPSOno® Lite (Transcription) so that as many as possible variants are generated to cover all possible cases a personal name can be transcribed. Do not use vocalized name with this feature set, the program will remove all diacritics automatically so no point in adding them.

Hazim	حازم

CAUTON: Use the fuzzy vocalization feature on your own risk, no guarantee of any kind that you will always get the correct vocalized version of the Arabic geographical name.

Partial vocalization

This feature allows auto-completion of the semi-vocalized names, based on the rules of Arabic spelling and is intended for partially vocalized names, the user have to add one or two diacritics enough to disambiguate the name in under the process of transcription; this is the default feature used by the program, the following example shows how to partial vocalization works.

Al-Suhayli	السُهَيْلي <- السُهيلي

Manual vocalization

The program does not interfere in the vocalization process if you choose to leave the name as is with the exception of forcing the application of some necessary basic spelling profiles such as taking into account the pronunciation of "Sun letters" and "*tā marbūṭa*". You must use this option if the names are already vocalized, in the example below not that we added the "*fatḥa*" before "*tā marbūṭa*" for the purpose of illustration, but we could omit it since the program would add anyway.

Muamar Al-Qadhafi	مُعَمَّر القَذَّافِي

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

The impact of each of the three modes is detailed in the following examples:

Vocalization	Input	Transcription	Details
Fuzzy	مهيلم	Muhailam	*Fa'tha* is added over *Laam*, generates 2 more orthographic variants
	مهيلم	Muhailim	*Kasra* is added below *Laam*, generates 2 more orthographic variants
	منهل	Manhal	*Fatha* is added over *Miim*, generates 6 more orthographic variants
	السليمي	As-Saliimi	medial and final *Yaa'* are considered long vowels , generates 3 more orthographic variants
Manual			
	السودان	As-Swdan	no change, *Siin* is dealt with as being *Sun letter*
	السَودان	As-Sawdan	*Fatha* over *Siin* is now considered
	الفَزّاني	Al-Fazzani	no change
	حسن	Hsn	no change
	البرغوثي	Al-Brghuthi	both *Yaa'* and *Waw* are considered a long vowels
Partial	رَفيدة	Rafiida	*Fatha Yaa'* is considered a long vowel
	رُفيدة	Rufaida	*Damma* is considered, *Sukuun* is added to *Yaa'*
	السوداني	As-Sudani	*Waw* is considered a long vowel
	السَوداني	As-Sawdani	*Sukuun* is added over *Waw*

MAPSOno® Lite (Transcription) undertakes the process of fuzzy vocalization based on statistical rules as well as applying some morphonemic criteria e.g. using Arabic morphological templates "*Binyanim*", even though the results may not come as expected. The professional edition (MAPSOno® Pro) is capable of doing full analysis of the personal name including the methods and tactics mentioned above so the results are much better.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Custom vocalization

MAPSOno® Lite (Transcription) allows other custom vocalization for the Arabic personal name, these are in fact primitive techniques based on direct matching with entries kept database either internal or external, the Lite version uses the internal one only:

Indirect vocalization

In addition to the vocalization modes discussed in the preceding paragraphs, the program provides direct and indirect vocalization techniques to improve the input Arabic personal names before the transcription process.

Vocalizing a name using built-in database

This edition allows the correction for a limited set of borrowed names and some common names of non Arabic origins e.g. (Abraham, Job, Ishmael, Jasmine) by direct matching and replacement using an integrated database which is also include some common titles i.e. honorific titles, check the box [*Database lookup*] in the Quick Options pane, the program will replace any occurrence of those names regardless of their current vocalization, this is shown in the following example:

Yasmin Esmaiil	ياسمين اسماعيل

If you want to use the names without the intervention of the program you must uncheck the box to disable using the integrated database, the name will then be transcribed the way you input.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Error detection and indication strategy

MAPSOno® Lite (Transcription) detects many types of errors in your input such as typos and inconsistent vocalization, for such errors the program will generally transcribe the name as usual and issue hints indicating those errors, the example below show how this works.

	Othmuan	*Inconsistent vocalization*	عُثمُان
	Al-Jaqri	*non Arabic pattern*	الجَقري

If the error is cannot be ignored such as insertion of strange symbols or Latin letters, the program will ignore the name completely and indicates the type of error. If you decide to transcribe the name, you should check the [*Bypass error check*] box from the Quick Options pane, you'll see the icon ( ) denoting that the error is bypassed; most of errors can be bypassed this way.

	Abdur-Rahman	عبدالرحمن

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Orthographic variants

MAPSOno® Lite (Transcription) generates variants for Arabic personal names according to the following two facts:

1. Orthographic changes that occur while adding the short vowels to the bare unvocalized name, this is a result of using the Fuzzy Vocalization feature.
2. Phonemic transcription which is, basically, some kind of mapping between the source phoneme and the closest matching phoneme in the target language (if any), this of course depends on how close the sound systems of both the source and target languages. The transcription process done by MAPSOno® Lite (Transcription) ,in this concern, incorporates many other factors as well.

The Fuzzy Vocalization technique generates variants using common Arabic name templates, thus the following name string (محمد) will give roughly six possibilities, shown below:

	Template	Template A	Template B
1	A[CuCaCCaC], B[CuCaCCiC]	مُحَمَّد /muħammad/	مُحَمِّد /muħammid/
2	A[CuCCaC], B[CuCCiC]	مُحْمَد /muħmad/	مُحْمِد /muħmid/
3	A[CaCCaC], B[CaCCiC]	مَحْمَد /maħmad/	مَحْمِد /maħmid/

All the above are legitimate forms both from orthographic and semantic point of view except for (3B) (/maħmid/) which is semantically meaningless. Orthographic variants generated this way do actually exceed few cases.

Phonemic transcription account for the large number of the generated variants, the number depends on the target language, regional variations of the source language, transcription precision.

Variant (1A) "/muħammad/" for example will generate at least three phonemic variants in both English and French languages.

Name	English	French
مُحَمَّد	Muhammad	Mouhammed
	Muhamad	Mouhamed
	Mohammad	Muhamed

Generated variants can be controlled in many ways:

- Eliminating them be checking the [*Variant suppression*] box in the Quick Options pane.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

- Limiting the number displayed from the Fine Tuning Interface, set the dial tagged [*Variants*] to the preferred value; setting the value to zero has the same effect as suppressing variants from the Quick Options pane.
- Deciding on how precise they should be, from the Fine Tuning Interface, set the dial tagged [*Precision*] to the preferred value; setting the value to 100% has the same effect as suppressing variants from the Quick Options pane.
- Variants can also be controlled by setting the Arabic language formality slider but this has minimal effect.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Displaying variants

MAPSOno® Lite (Transcription) doesn't display all variants by default, click on the (+) before the variant icon (☰) to expand the name variants, you may expand the whole variants tree by clicking (**Expand all**) button, click again to collapse the tree; information about the generated variants can be found in three different columns illustrated below:

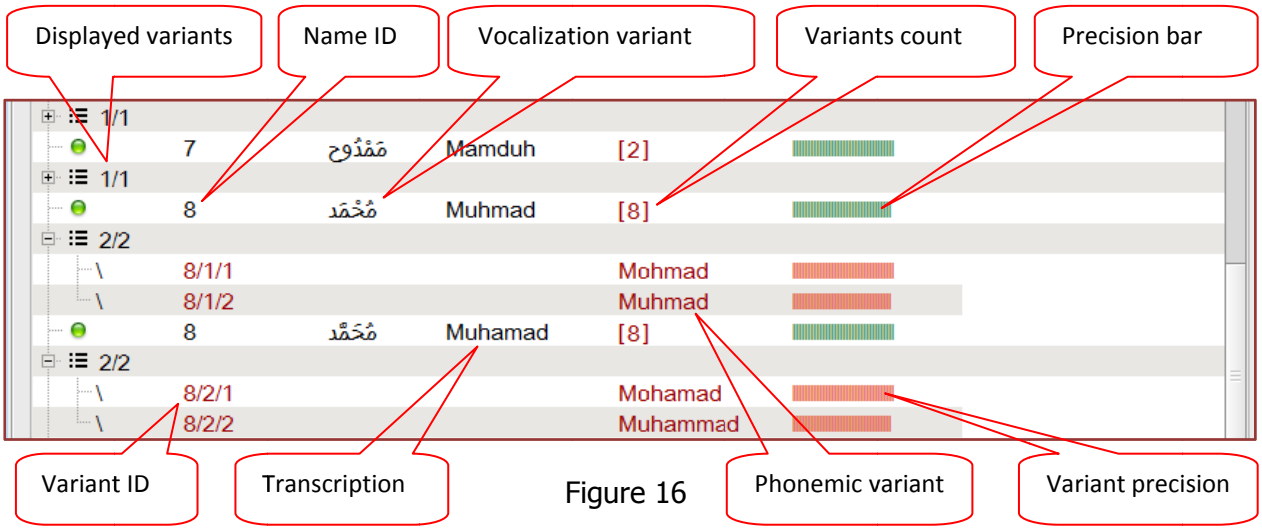


Figure 16

Column	Function	Details
Icons	currently displayed variants	This icon (☰) means the name has variants, the number configured as A/B, where A is the top most precise variants while B is the total number of variants that passed the precision criteria.
ID	variant ID	Variant ID is configured as A/B/C where A is the same as the original name ID; B is orthographic variants serial number i.e. the order of the Arabic variant generated using the Fuzzy Vocalization feature. C indicates the serial number of the phonemic variant generated for every orthographic variant; this number is reset each time a new orthographic variant is generated.
Variants	variant and total number of variants	Variants are sorted in a descending order beginning with the most precise one; chances are that you may discover a variant that is more precise than the main entry transcription. The total number of the *generated* variants is shown between the red brackets in the same line of the original input name. The only way to tell whether the name has more variants is by referring to this number which indicates the maximum number of actually generated (if any) variants regardless of the criteria set in the Fine Tuning interface.

Variants have no assigned status icons, they follow the original name status; each variant has a precision bar indicator.

Variant is not prepended with an ID prefix; this will programmatically ease the post processing.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Precision calculation

MAPSOno® Lite (Transcription) calculates transcription precision based on many criteria applied two levels, the first based merely on some shallow empirical orthographic matching factors, the second uses complicated formula to calculate the phonetic distance between the whole transcribed name and the *original* sound in Arabic language thus showing how similar or close they are; this can be ranked based on their similarity or difference score which is eventually shown graphically as percentage bars. No precision is calculated for names with fatal errors.

Precision bars

Shown in the sixth column are the precision indicator bars, each line represents 2% of the calculated transcription precision; hover over the bars to see the numerical values, these numerical values will be printed and saved instead of the bars.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Preparing for output

Printing

The program prints the transcribed names as displayed in the View Interface except for the precision bars, status icons will be printed for the purposes of proofreading, make sure to make necessary setup before printing.

If you would like to carry out printing using some other desktop publishing applications e.g. Microsoft Word then you have to save the result to a file in the (RTF) format because symbols will be dealt with internally in order to display correctly on third party software, the (RTF) format ensure correct display and print without the need for further processing.

If you want to save the names for subsequent processing then you have to save the files using the (UTF8) encoding since almost all commercial software can deal with text files encoded in this global encoding otherwise you may lose some data or be forced to install additional fonts; example of such third party software supporting UTF8:

> Microsoft Word®, Microsoft Access®, Microsoft Excel®, Microsoft FoxPro®, Notes®, Open Office®

Publishing on web requires saving the files in (HTML) format, the program supports the standard HTML therefore we do not expect the result to look different in the other html browsers, we already tested the following software products which displayed the same looking of MAPSOno® Lite (Transcription) HTML formatted output files:

> Microsoft Internet Explorer®, Mozilla Firefox®, Google Chrome®, Netscape Navigator®, Opera®

Saving files in different formats and encodings

Output files can be saved in more than one format and encoding as long as the file is open, select the format and then click (**Save**) button; you cannot save the file in more than one transcription language.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Output sorting and filtering

Alphabetical sorting is already described in paragraph (Result display area) on page (18).

To filter the transcription results click on the master function (**Filters**) button, a panel will appear beneath the display area, as shown in (Figure 17) below, you must specify the filter criteria to display only those names matching the criteria you specified, available criteria are listed in the table below.

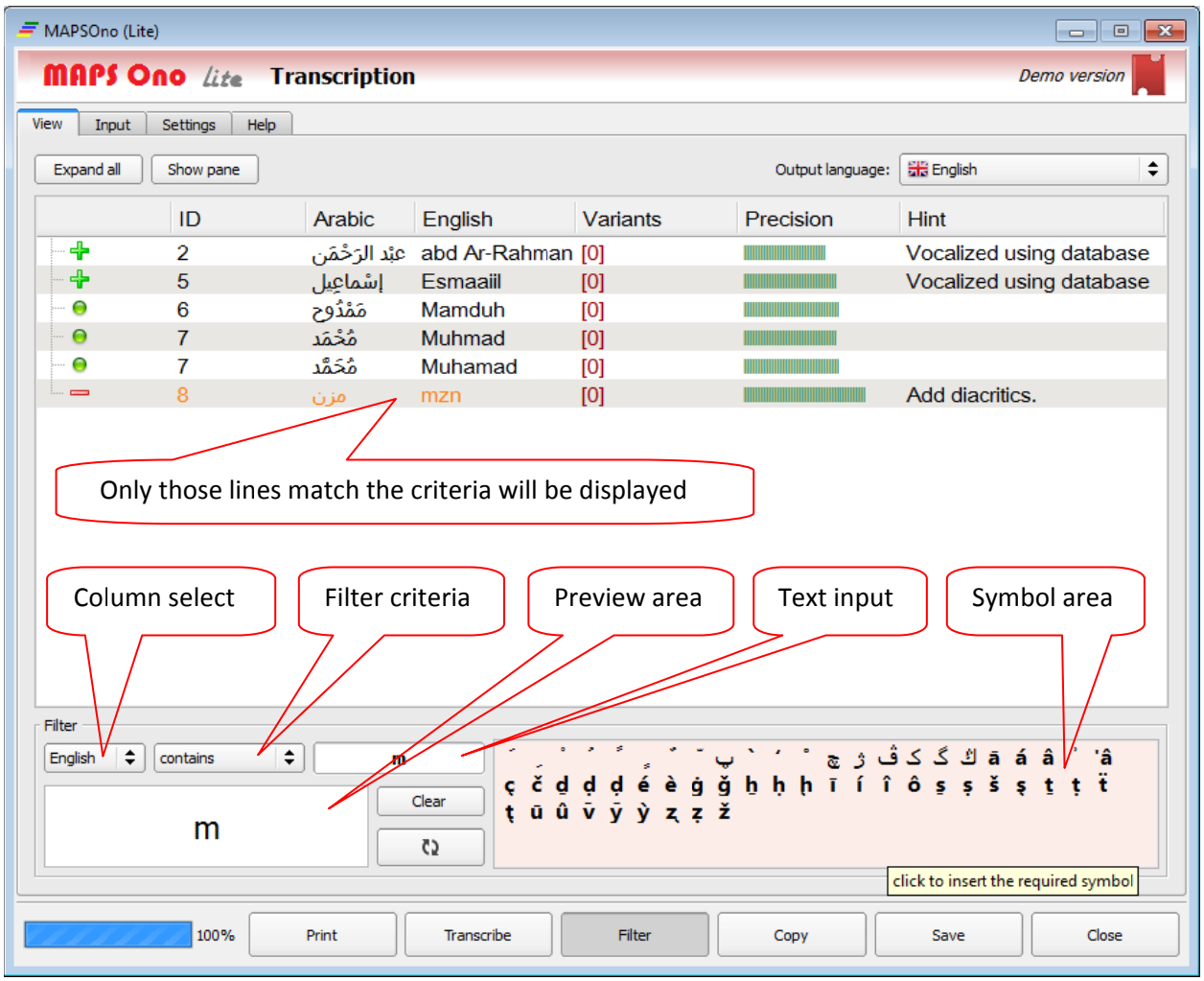


Figure 17

Criteria	Details
Contains	To display the lines contain a specific text or symbols
Does not contain	To display the lines that do contain a specific text or symbols
Begins with	To display the lines begin with a specific text or symbol
Ends with	To display the lines end with a specific text or symbol

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

The filter panel displays some symbols, two dropdown menus, text input box, in addition to a filtered sample preview area; all are described as follows:

- Symbols area: click on the symbols to insert it in the text input box, this serves as a virtual keyboard if the specified symbols is not available on the keyboard.
- Column selection dropdown menu: select the column on which you like perform the filtering; the whole result contents will be affected.
- Criteria selection dropdown menu: select the criteria you want to apply on the results.
- Text input box: enter the text or any part of text here.

While you are setting the filtering criteria a real data sample will be displayed in the preview area, when you done setting the filtering criteria, click on (**OK**) button to start filtering the results, only the matching lines will be displayed as shown in on (Figure 16) previous page.

Your filter criteria will be displayed in the preview area, when you done setting the filtering criteria, click on the ( ) button to start filtering the results, only the matching lines will be displayed.

Note: please be informed that filtering is case sensitive.

After filtering the results the content of the display area will change completely, the only way to restore the original results is delete the filter criteria text and click the ( ) button again.

If the (**Filter**) button is showing the text (Filter) then the results are original, the case is different after filtering where the same button will show the icon ( ) instead.

Filtered out results will still be saved and printed unless you intentionally set the maximum length of the input names or limit the generated variants.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Transcription scenarios

The following are examples of how to transcribe Arabic personal names for different purposes and situations that you may encounter, this does not necessarily mean that you should abide by the techniques followed in these examples, consult the help of the client asked for the transcription services to provide information about the specifications on a case by case basis.

General directions

The fundamental issues that need to be addressed summarized below:

1. Arabic personal names input file encoding.
2. Arabic personal names input file format.
3. Name of the transcription language.
4. Required output file format.
5. Required output file encoding.
6. Required precision of variants.
7. Whether the client is going to make subsequent processing.

A. Transcribing names for an official foreign client

A foreign official organization asked for transcription of certain Arabic personal names, this organization is not authorized to reuse of data (e.g. modify or sell) but it can publish on the Internet in a format that does not allow retrieval i.e. copying or printing; use the following settings:

1. Transcription language: (decided by the client)
2. Input file encoding: [Windows-1256 or UTF-8]
3. Output file encoding: [UTF-8]
4. Output file format: [RTF]
5. Open the result file in a text formatter e.g. Microsoft Word, and use Tahoma font size 16, for Arabic text and Arial Unicode size 12 for the target language script, you can also use the table auto format feature available in Microsoft Word; you may format the data using Microsoft Excel.
6. Convert the file to PDF format and add the appropriate protection (disable copying, editing, and printing).

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

B. Transcribing names for a foreign client

A foreign non-Arabic speaking official organization asked for transcription of certain Arabic personal names, this organization is authorized to reuse of data (i.e. test, modify); use the following settings:

1. Transcription language: (decided by the client)
2. Input file encoding: [Windows-1256 or UTF-8]
3. Output file encoding: [UTF-8]
4. Output file format: [CSV or Tab delimited]
5. Transliteration of output file: [KATS] or [SATTS] or [Buckwalter]
6. (KATS) transliteration system will convert Arabic text to English ASCII text which can be converted back to Arabic script without losing any data.
7. Send the file together with a copy of the (KATS) transliteration system which is available at Kalmasoft website here: http://www.kalmasoft.com/downloads/KATS.zip.

C. Transcribing names for an official home client

A home organization asked for transcription of certain personal names to later be used by some official internal body e.g. national archive, competent ministry, technical office, data management division etc., this organization is authorized to reuse of data (i.e. publish, modify), use the following settings:

1. Transcription language: (decided by the client)
2. Input file encoding: [Windows-1256 or UTF-8]
3. Output file encoding: [UTF-8]
4. Output file format: [CSV, SQL, or Tab delimited]
5. The settings above allow opening the output in almost all database management software and spreadsheets such as (Oracle, Microsoft Access, Microsoft FoxPro, MySQL, Microsoft Excel, Microsoft SQL Server, Lotus Notes).

Using the included sample files

The (Samples) folder contains files of Arabic personal names from all states in the Arab region, you can use these files directly for testing, use Microsoft WordPad to open each file to view or edit the contents, load the file and start transcription, you may copy and paste to the input pane if you like. The professional edition allows loading all these files at once without having to open each file separately.

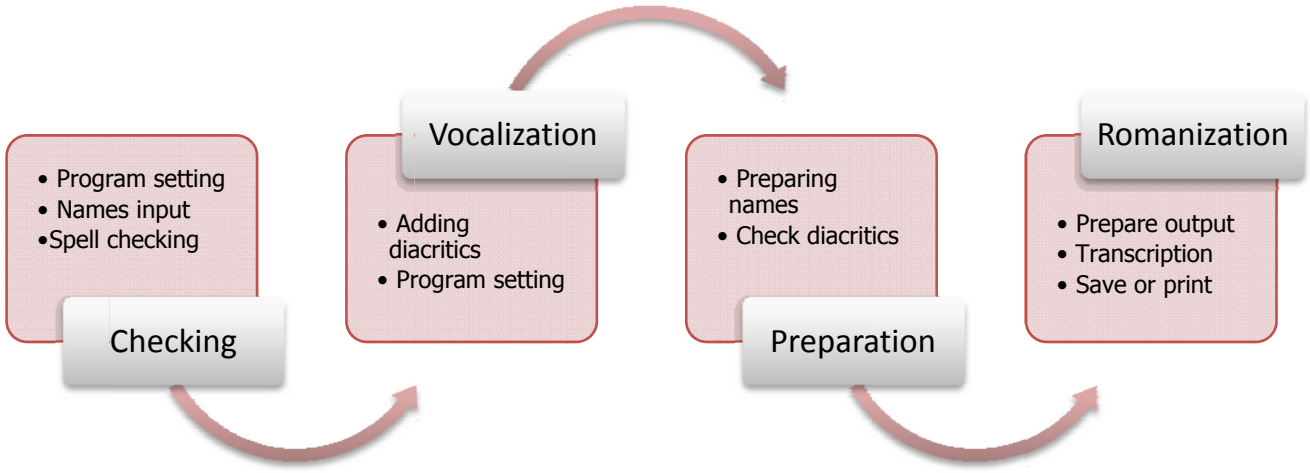
Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Tips and tricks

The following are few notes to speed up the work and add up for a professional using of the program:

1. You need not have to click the buttons to determine the input source of personal names (single line, input pane, load file), this will be done automatically for you once you start typing, pasting, or loading files, into any of the input ports of the program.
2. You do not have to click the (**Transcribe**) button every time to start transcription; you can use the (**Preview**) button to perform the same function.
3. If you are not sure of the validity or correctness of the Arabic personal names you can select the [*Check mode*] from the master language selector menu, this mode provides a basic spell check for the input names, store the list of names and edit or delete unwanted ones.
4. Click on the header of any column to sort its contents alphabetically, click again to reverse the sorting order.
5. If your data takes time to prepare e.g. by adding diacritics, the best way is to save your favorite settings and then run the program on it later.
6. To adjust the width of any column in the View Interface, you do not need to drag its borders each time, double-click between the columns headers and it will be adjusted to accommodate the contents according to the longest item in it.
7. Beware that Fuzzy Vocalization generates parallel vocalized Arabic names, your original input is not echoed or even saved, [*Apply Fuzzy Vocalization*] works fine for legitimate Arabic personal names.
8. A best way to transcribe a list of name is to start with [*Apply Fuzzy Vocalization*], vocalized input can then be input and transcribed using the [*Partial Vocalization*]; you may try to use the [*Use input name as is*] option just to make sure you get it right.

If you use the program for the first time, we recommend you proceed with Transcription according to the illustration below:



Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

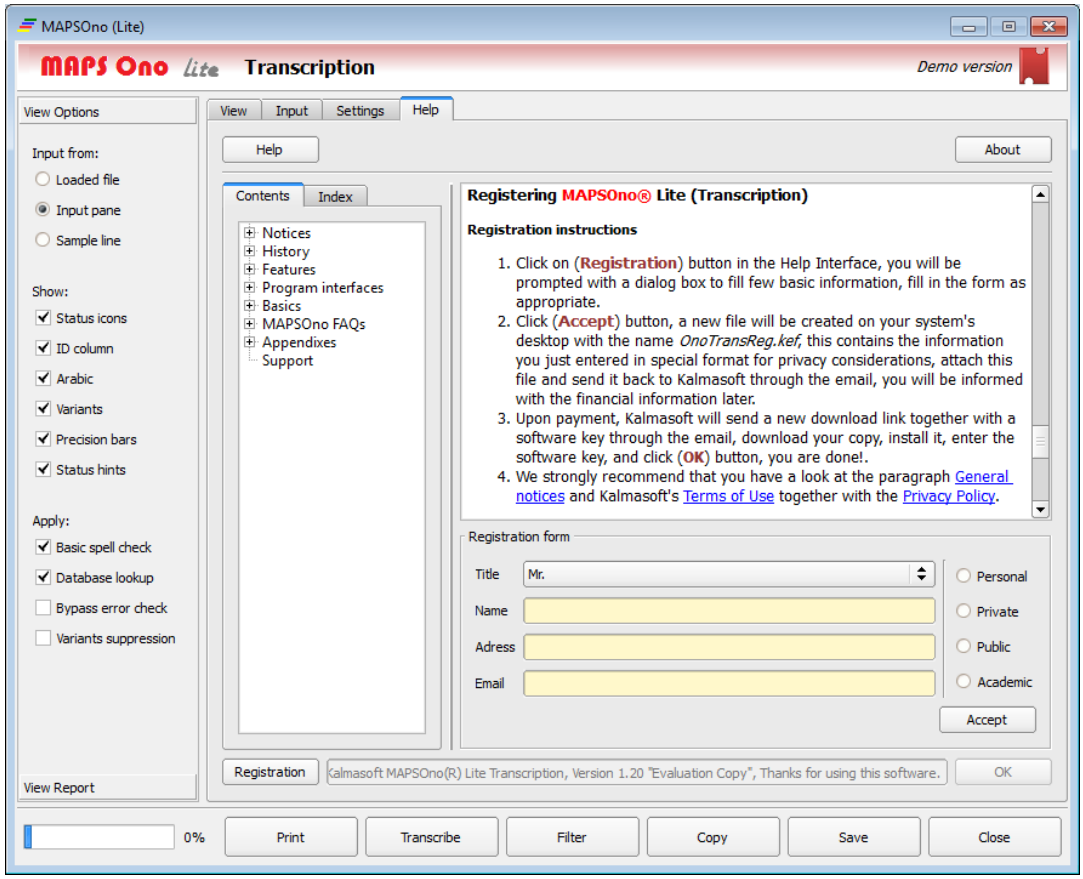
MAPSOno® frequently asked questions

Your problem may already be solved here:

Registering MAPSOno® Lite (Transcription)

Registration instructions

1. Click on (**Registration**) button in the Help Interface, you will be prompted with a dialog box to fill few basic information, fill in the form as appropriate.
2. Click (**Accept**) button, a new file will be created on your system's desktop with the name *OnoTransReg.kef*, this contains the information you just entered in special format for privacy considerations, attach this file and send it back to Kalmasoft through the email, you will be informed with the financial information later.
3. Upon payment, Kalmasoft will send a new download link together with a software key through the email, download your copy, install it, enter the software key, and click (**OK**) button, you are done!.
4. We strongly recommend that you have a look at the usage policy in page (7) and Kalmasoft Terms of Use http://www.kalmasoft.com/terms.htm together with the Privacy Policy http://www.kalmasoft.com/privacy.htm .



Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Input interface issues

Can't load files

- MAPSOno® Lite (Transcription) is not suited for ordinary running text, if you are trying to load files containing long paragraphs this will not work.
- Make sure that the file does not contain any strange symbols, you may comment out unwanted lines by adding "*" or "#"; those lines will appear again in the header part of the saved or printed output.
- You may be trying to load an unsupported file type, please check the list of supported types in paragraph Input file format and encoding pane in page (25), if your file is not supported the easiest way is to copy the names and paste them to the Input pane please refer to page (17); you may also save the file as txt from the native application.

Program halts during load

- You are trying to load a huge file that could possibly make the program irresponsive, please interrupt the program, divide the file, and try again.

Setting issues

My settings don't work

- Some settings may only be applicable under certain circumstances, for instance, the [*Database lookup*] is useless if no database available.
- Some settings may look contradicting for the first time, this not the case, however, if you think that you have messed things up please close the program and start again; if you have already saved any custom settings please restore the defaults and try again with a fresh settings, this is described in paragraph Settings interface in page (22).

My new settings do not seem to affect the results

- Some specific settings are not available in the trial version e.g. language formality and regional dialects for both source and target languages, these have been intentionally disabled.
- Some settings are sensitive only to specific names with irregular patterns e.g. *Alif Maqsura* and *Taa' Marbuta*, so it is not necessarily that they affect other regular names.

Can't save my settings

- This feature is not available in the trial version; new settings are active only during the transcription session. The program will not leave "ini" or "cfg" files in your system.
- Register the software to get full access to this feature, please refer to page (52).

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Display issues

Interface language sounds "Greek"

- Make sure that you select the right language for your interface, if it still looks strange check if you have the right fonts installed in your system.
- MAPSOno® Lite (Transcription) trial version support only English and Arabic for the interface, supported target languages include (Amharic, English, French, Japanese, and Russian).

Transcribed names not showing up

- Please make sure you select the right input source, nothing will be displayed on the View Interface if you choose to input names from the single line.
- If you are loading names from a file please check for any strange symbols preceding the name e.g. (*, #).

Precision indicator bars shown scattered

- Install the (Arial Unicode MS) TrueType font and try again.

Precision bars never hit 100%

- They should not! never expect over 90% at the best, remember that "transcription" is about migrating phonological characteristics between languages with sound systems that may actually be so different, hence best hits can be made for languages with their sound systems being as close as possible to Arabic; those are definitely the Semitic family languages i.e. Amharic, Tigrinya and Hebrew, the poorest precision comes with Asian languages, remarkably Chinese.

I am getting too many variants

- As a rule of thumb, less variants means better phonetic matching between the Arabic name and its transcripts; Well, here are some efficient ways to keep the variants down:
    1. Kill variants! or set the max top precise ones to minimum that satisfies your needs, you don't need hundred variants generated for every Arabic name you input do you?
    2. Shorten the names, not by setting the clipping limit to the minimum, this has no effect and will not work, get a list of unique Arabic names instead of loading full names.

Can't copy items

- This feature is not available in the trial version; you cannot copy items from the View Interface.
- Register the software to get full access to this feature, please refer to page (52).

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Print issues

Printout not showing results

- This feature is not available in the trial version; only the original input and the transcription report can be printed.
- Register the software to get full access to this feature, please refer to page (52).

Mojibake and strange symbols when printing

- Please make sure that you select the right target language.
- If you are transcribing Arabic names to some Asian language please make sure that your system is already set properly to display the Asian characters.
- If the above fail, reinstall the (Arial Unicode MS) font and try again.

How can I get rid of Kalmasoft logo

- Kalmasoft logo "the four horizontally stacked color bars" is usually printed at the top of the first page, if you do not like this to be inserted in your document save the result to (RTF) file and remove the logo in a suitable text editor and put whatever you like.
- If you would to have your organization logo printed instead please contact Kalmasoft, direct contact details can be found here http://www.kalmasoft.com/contact.htm .

File saving issues

Can't save files

- By default MAPSOno® Lite (Transcription) saves result files in the (Input) folder, please set the new path clearly in the dialog box.
- Check the file name and file path.

Can't find my results in the saved file

- This feature is not available in the trial version; results can only be displayed.
- Register the software to get full access to this feature, please refer to page (52).

Filtering and sorting issues

Filtering doesn't work

- Check the criteria you are trying to apply.
- Make sure that the results are not filtered already, clear the current filter text and just click ( ) button to restore the original results.
- Beware that filtering is case sensitive.
- Perhaps the results do not actually contain the filter text.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Can't sort results

- This feature is not available in the trial version; save the result to a file in a suitable format and do sorting using other tools.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Appendixes

Appendix A. Error messages and status hints

Messages and tips do not necessarily always indicate errors, some describe the status after you perform certain operations; some icons in the table below may not be available in the edition or the version of MAPSOno® Lite that you use.

Icon	Message	Hint	Meaning
	Not a name, ignored.	Check for Latin	Multilingual contents is not allowed
	Name is too long, clipped.	Name exceeding the maximum	Total length should not exceed 25 characters, default is 10
	Done		
	Strange contents.	The name has strange contents	e.g. "?", "+", "#", "@"
	Special letters	The name has special letters	e.g. . "چ", "ڤ", "ڷ"
	Unable to vocalize		Need no vocalization or ready vocalized
	Non-Arabic name		
	This name is *Nunated*	Remove *Tanwīn*	Names should not be *Nuunated*
	Repetition		Repeated content
	No diacritics	Vocalize the name	Vowels stripped out
	Extended letters		e.g. . "گ", "ۓ"
	Added from the database.		This name is added directly from the database i.e. not processed
	Spelling error		e.g. name starting with tā' marbūṭa
	Inconsistent vocalization	Check the vowels	e.g. adding *Ḍamma* before *tā marbūṭa*
	Variants exist.	Activate variant generator	Professional Edition only
	Re-enter the name		
	Name is too short, ignored		The minimum length is 3 characters, spaces and diacritics counted

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

The table below shows the status-only icons:

Icon	Message	Hint	Meaning
	Correct		Linguistically correct
	Apparently correct	Recheck the name	Linguistically correct but recheck
	Corrected		This name has been corrected
	Marked for exclusion	Name to be excluded	e.g. exclusion from Transcription
	Check bypassed		Error checking has not been applied

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Appendix B. Romanization systems

System	transcription	Full name
ADEGN (*)	Transcription	Arabic Conference on Geographical Names
ALA-LC	Transcription	Association of American Libraries -Library of Congress
BGN	Transcription	Board of Geographical Names/ Permanent Committee on Geographical Names
Buckwalter	Transliteration	Tim Buckwalter Arabic Transliteration
DIN31635	Transcription	Deutsches Institut für Normung
IGN	Transcription	Institut géographique national
ISO233	Transcription	International Standards Organization
KATS	Transliteration	Kalmasoft Arabic Transliteration System http://www.kalmasoft.com/devtool.htm
RJGC	Transcription	Royal Jordanian Geographic Center
SAS	Transcription	Spanish Arabists School
SATTS	Transliteration	Standard Arabic Technical Transliteration System
UNGEGN	Transcription	United Nations Group of Experts on Geographical Names

(*) Not official name, another acronym is "ACGN", please consult the relative sources.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip

Glossary of terms

Term	Meaning
Fuzzy Vocalization	An automated way of adding short vocals to Arabic text based on morphological and statistical analysis.
KATS	Kalmasoft Arabic Transliteration System, please refer to Kalmasoft development tools page.
orthographic variants	Arabic names of common grapheme that spell different because of the lack of Arabic short vowels
vocalization	For Arabic, this means adding the short vowel diacritics; these have a major impact in disambiguating the possible ways a Arabic name is pronounced.

Link: http://www.kalmasoft.com/support/DOCMOLTrans.zip